JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
Vol. 43, No. 4 AMERICAN WATER RESOURCES ASSOCIATION August 2007

AUTOMATIC CALIBRATION OF HYDROLOGIC MODELS WITH MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM AND PARETO OPTIMIZATION[1]

Remegio B. Confesor Jr. and Gerald W. Whittaker[2]

ABSTRACT: In optimization problems with at least two conflicting objectives, a set of solutions rather than a unique one exists because of the trade-offs between these objectives. A Pareto optimal solution set is achieved when a solution cannot be improved upon without degrading at least one of its objective criteria. This study investigated the application of multi-objective evolutionary algorithm (MOEA) and Pareto ordering optimization in the automatic calibration of the Soil and Water Assessment Tool (SWAT), a process-based, semi-distributed, and continuous hydrologic model. The nondominated sorting genetic algorithm II (NSGA-II), a fast and recent MOEA, and SWAT were called in FORTRAN from a parallel genetic algorithm library (PGAPACK) to determine the Pareto optimal set. A total of 139 parameter values were simultaneously and explicitly optimized in the calibration. The calibrated SWAT model simulated well the daily streamflow of the Calapooia watershed for a 3-year period. The daily Nash-Sutcliffe coefficients were 0.86 at calibration and 0.81 at validation. Automatic multi-objective calibration of a complex watershed model was successfully implemented using Pareto ordering and MOEA. Future studies include simultaneous automatic calibration of water quality and quantity parameters and the application of Pareto optimization in decision and policy-making problems related to conflicting objectives of economics and environmental quality.

(KEY TERMS: simulation; optimization; nonpoint source pollution; Pareto; multi-objective evolutionary algorithm; automatic calibration.)

Confesor, Remegio B., Jr., and Gerald W. Whittaker, 2007. Automatic Calibration of Hydrologic Models With Multi-Objective Evolutionary Algorithm and Pareto Optimization. *Journal of the American Water Resources Association* (JAWRA) 43(4):981–989. DOI: 10.1111/j.1752-1688.2007.00080.x

INTRODUCTION

Most hydrologic models are characterized by complex functional relationships and a large number of parameters that are usually conceptual representations of the watershed. In most cases, these parameters cannot be directly measured or are not exactly known because of spatial variability and measurement error, hence the need for model calibration (Gupta *et al.*, 1998; Eckhardt and Arnold, 2001). Calibration is the estimation of model parameters to achieve a system that closely resembles the actual system that the model represents (Yapo *et al.*, 1998). Manual trial and error calibration is time consuming and depends on the modeler's experience, skill, and knowledge of the model's processes and dynamics. While overcoming the shortcomings of manual trial

[1]Paper No. J06024 of the *Journal of the American Water Resources Association* (JAWRA). Received February 10, 2006; accepted November 16, 2006. **Discussions are open until February 1, 2008**.

[2]Respectively, Research Associate, Department of Agricultural and Resource Economics, 213 Ballard Extension Hall, Oregon State University, Corvallis, Oregon 97331; and Research Hydrologist, USDA-ARS, National Forage Seed Production Research Center, 3450 SW Campus Way, Corvallis, Oregon 97331 (E-Mail/Confesor: confesor@onid.orst.edu).

and error calibration, automatic calibration of complex hydrologic models requires a tremendous amount of computation.

Multiple objective optimizations improve model calibration but further increase the computational requirements. Classical multi-objective optimization involves transforming multiple objectives into a single function. The most common method is the weighted sum principle where the objectives are multiplied with user-defined weights and added together to form a single function. To avoid the disadvantages of converting multi-objective functions into a single optimization problem, Pareto optimization has gained use and popularity in hydrologic modeling. In problems with at least two conflicting objectives, a set of optimal solutions exists as a result of the trade-offs between these objectives. A Pareto optimal solution set is achieved when a solution cannot be improved upon without degrading at least one of its objective criteria. The cardinality of the Pareto optimal set is one if the objective functions are not conflicting to each other (Deb, 2001). This means that the optimum solution corresponding to any objective is the same and the Pareto front in the search space will converge into a single solution. Otherwise, a front of different solutions is searched for conflicting objectives. This optimal front can be established after several iterations and function evaluations that require tremendous amount of computation. Population evolution-based search algorithms such as genetic algorithms (GAs) (Holland, 1975; Goldberg, 1989) and the shuffled complex evolution (SCE) algorithm (Duan *et al.*, 1992) have been used in implementing Pareto ranking.

Yapo *et al.* (1998) and Vrugt *et al.* (2003), while calibrating the conceptual rainfall-runoff model Sacramento Soil Moisture Accounting (SAC-SMA), incorporated dominance or Pareto ranking into the SCE so that the population evolved toward the Pareto optimal set in the search space. Eckhardt and Arnold (2001) also used the SCE and took 6 days to calibrate a distributed catchment hydrologic model that explicitly optimized 18 parameters and simultaneously adjusted 143 parameters in fixed ratios. Khu and Madsen (2005) used a modern and fast nondominated sorting GA in a multiple objective automatic calibration of another conceptual rainfall-runoff model (MIKE11/NAM). However, all of the above calibrations were burden by the computational requirement imposed by the GA or SCE in exploring the entire feasible search space for the Pareto optimal set. In calibrating a process-based and semi-distributed hydrologic model, van Griensven and Bauwens (2003) normalized and put weights in their objective functions to create a global optimization criterion (GOC) with the SCE, hence overcoming the computational burden.

This study used a Beowulf cluster consisting of a server (P4 3.2 GHz dual processor) and 12 computation nodes (P4 2.4 GHz) at the National Forage Seed Production Research Center (NFSPRC), a unit of the U.S. Department of Agriculture – Agricultural Research Service (USDA-ARS) in Corvallis, Oregon (Whittaker, 2004). With the NFSPRC Beowulf cluster providing the computational power required in Pareto ranking and multi-objective evolutionary algorithm (MOEA), the application of Pareto ordering optimization in the multiple-objective automatic calibration of a complex process-based, semi-distributed, and continuous hydrologic model such as the Soil and Water Assessment Tool (SWAT) was investigated. This study also demonstrated the generation and advantage of Pareto solutions in optimizing selected SWAT parameters with two objective functions.

METHODS

NSGA-II and PGAPACK

Deb *et al.* (2002) proposed the nondominated sorting genetic algorithm II (NSGA-II), a fast and efficient MOEA characterized by a nondominated sorting algorithm, an elitist selection method, and the elimination of a sharing parameter. NSGA-II assigns fitness by Pareto ranking (or nondomination) and crowding distance to the combined parent and child populations. The solution is then ranked according to the number of solutions that dominates it. A solution X_1 dominates another solution X_2 if both conditions are satisfied (Deb, 2001) (1) the solution X_1 is no worse than X_2 in all objectives; and (2) the solution X_1 is strictly better than X_2 in at least one objective.

Crowding distance is the average distance between an individual and its nearest neighbors in the search space (see Deb *et al.*, 2002). In minimization problems, solutions that are dominated by fewer solutions (i.e., has a lower rank) are given a better fitness than the dominated ones. In cases where the solutions have the same nondomination rank, the solution with larger crowding distance is preferred, thus ensuring diverse and well-spread population. The new parent population is chosen from the combined parent and child population based on the solutions' fitness or rank, thus the elitist selection. NSGA-II had been tested and yielded solutions that converged to the true Pareto front of problems with convex, nonconvex, nonconvex disconnected, convex disconnected, and nonconvex nonuniformly distributed solutions (Deb *et al.*, 2002). The NSGA-II code is available from the Kanpur Genetic Algorithms Laboratory at http://www.iitk.ac.in/kangal/codes.shtml.

PGAPACK is a general purpose, data-structure-neutral, parallel GA library developed at the Argonne National Laboratory (Levine, 1996). Its key features include: (1) parallel portability across uni-processors, multiprocessors, multi-computers, and workstation networks; (2) callable in C and FORTRAN languages; (3) binary-, integer-, real-, and character-valued native data types; (4) simple interface for novice and expert application users; (5) large set of example problems; and (6) parameterized population replacement. The structure and usage of PGAPACK is discussed in detail in Levine (1996).

The SWAT Model

The Soil and Water Assessment Tool (SWAT) (Arnold *et al.*, 1998) was developed by the (USDA-ARS) "to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over long periods of time" (Neitsch *et al.*, 2002, p. 1). SWAT is physically based, uses readily available inputs, computationally efficient, and is a continuous model that operates on a daily time step (Neitsch *et al.*, 2002). SWAT is not designed to simulate single-event storms. In SWAT, the entire watershed can be divided into several subbasins and each subbasin is further divided into unique combinations of land use and soil properties called the hydrologic response unit (HRU). However, the location of each HRU is not specified in the subbasin. A graphical user Geographic Information System (GIS) interface (AVSWAT2000) can be used to input and designate land use, soil, weather, ground water, water use, management, pond and stream water quality data, and the simulation period (Di Luzio *et al.*, 2001). GIS input files include digital elevation model (DEM), land use and soil properties layers, and weather database.

The Calapooia River Watershed

The Calapooia river watershed (U.S. Geological Survey, USGS, 10 digit HUC 1709000303) is a tributary of the Willamette River basin west of the Cascades mountain range in Oregon (Figure 1). It has drainage area of 963 km^2 as delineated from a USGS streamflow gaging station (44°37′15″N, 123°07′40″W) in Albany, Oregon. Its elevation ranged from 56 to 1,576 m and its land use is mainly agriculture (43%), forest (41.8%), and hay/pasture/range areas (11.2%). The remaining areas were composed of wetlands, urban areas, and water bodies. The watershed has a predominantly winter rainfall climate with precipitation from December through February comprising about 50% of the

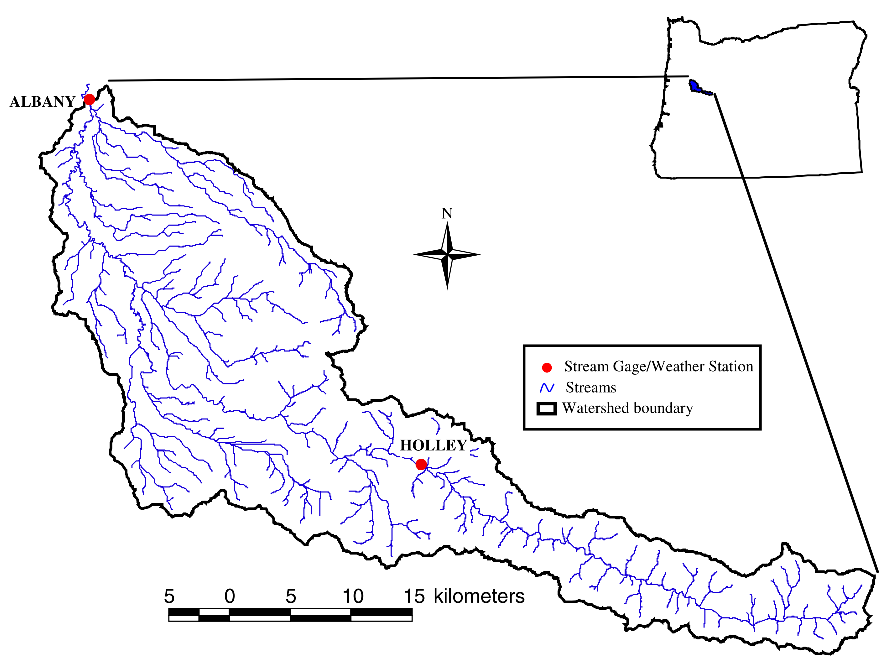

FIGURE 1. The Calapooia River Watershed in Oregon, USA.

annual total with lesser amounts in the spring and fall and negligible precipitation during summer. The precipitation, minimum and maximum temperature data were taken from the weather station at Holley, Oregon (44°21′5″N, 122°47′10″W, elevation = 165 m).

The watershed was delineated with a threshold size of 150 ha using the Arcview interface to SWAT2000 (Di Luzio *et al.*, 2001). The 10-m DEM used in delineating the watershed was taken from the Regional Ecosystem Office (http://www.reo.gov/reo/data/DEM_Files/indexes/orequadindex.asp). The observed daily streamflow data used in calibrating SWAT were obtained from the USGS National Water Information System (NWIS) website (http://nwis.waterdata.usgs.gov/or/nwis/sw). The state soil geographic (STATSGO) database for Oregon was from the U.S. Department of Agriculture – National Resources Conservation Service, USDA-NCRS (http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo). Land use for the Willamette basin was acquired from the USGS National Water-Quality Assessment (NAWQA) Program (http://or.water.usgs.gov/projs_dir/pn366/landuse.html). Climate data were taken from the Oregon Climatic Service (http://www.ocs.oregonstate.edu/).

Automatic Calibration

The SWAT model was initially set up using the Arcview interface (AVSWAT2000) to SWAT (Di Luzio *et al.*, 2001). HRU distribution was defined by eliminating landuses that were <10% of the watershed area and then by removing soil types that were <10% within a land-use area. These thresholds resulted in four dominant land uses (mixed forest, evergreen forest, perennial grass, and hay, pasture, and rangelands) and nine major soil groups in the watershed. After overlaying the landuse and soil properties layers, there were a total of 17 HRUs. Based on the SWAT user's manual (Neitsch *et al.*, 2002) and previous studies (Eckhardt and Arnold, 2001; Van Liew *et al.*, 2005), 11 parameters (eight for each HRU and three for the whole watershed) were used in the calibration, resulting in 139 unique parameter values to be explicitly optimized. The limits of the parameters for calibration were fixed (Table 1) to ensure realistic and acceptable values representative of the watershed characteristics. The calibration (October 1973 to September 1976) and validation (October 1976 to September 1979) periods were set for three water years due to the availability of streamflow data.

PGAPACK was called in FORTRAN to randomly generate the initial parent population of 100 solutions. A child population (size = 100) was then generated through selection, mutation, and crossover from the parent population with PGAPACK (Figure 2). In the first iteration, the child and parent populations were evaluated for two objective functions (see next section). In the evaluation step, SWAT was called as a subroutine and was executed for each solution.

Objective Functions and Model Evaluation

The objective functions were to minimize the average root mean square error (RMSE) of the observed *vs.* simulated peak (driven) flows and to minimize the average RMSE of the observed *vs.*

TABLE 1. Range of Values of Parameters Selected for the Calibration of the SWAT Model.[a]

Variable	Description	Minimum	Maximum
GWDELAY	Ground-water delay time (days)	0.001	62.000
ALPHABF	Base-flow alpha factor (days)	0.040	1.000
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	0.000	100.000
GWREVAP	Ground-water "revap" coefficient	0.020	0.750
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur (mm)	0.000	200.000
AWHC[b]	Available water capacity of the upper-most soil layer (mm H_2O/mm-soil)	0.110	0.300
ESCO	Soil evaporation compensation factor	0.700	1.000
CN2[b]	SCS runoff curve number	55.000	89.000
CHK2[c]	Effective hydraulic conductivity in the main channel alluvium (mm/h)	0.000	150.000
CHK1[c]	Effective hydraulic conductivity in tributary channel alluvium (mm/h)	0.000	150.000
SURLAG[c]	Surface runoff lag coefficient	1.000	21.000

[a]Unless specified, the range of values is the same for all the 17 hydrologic response units (HRUs).
[b]For brevity, only the range of values across all HRUs is shown. Each HRU has a different range depending on soil type or land use.
[c]Only one value of this variable is calculated for the whole watershed.

simulated low (nondriven) flows. The RMSE was defined as

$$\mathrm{RMSE} = \left[\frac{1}{n}\sum_{i=1}^{n}\left(Q_{\mathrm{obs},i} - Q_{\mathrm{sim},i}\right)^2\right]^{0.5}, \tag{1}$$

where n is the number of time steps with peak or low flow events, $Q_{\mathrm{obs},i}$ is the observed streamflow at time i, and $Q_{\mathrm{sim},i}$ is the simulated streamflow at time i.

The hydrographs were partitioned into driven and nondriven components assuming that the behavior of the watershed is different during the periods driven by rainfall and periods without rain (Boyle *et al.*, 2001). The periods immediately after rainfall (nondriven quick) should be dominated by interflow and the latter periods (nondriven slow) by base flow. The driven flow can be associated with the rising limb of the hydrograph and the nondriven flow with the recession. A base-flow filter (Arnold *et al.*, 1995; Arnold and Allen, 1999) was used to estimate the base-flow component of the observed streamflow.

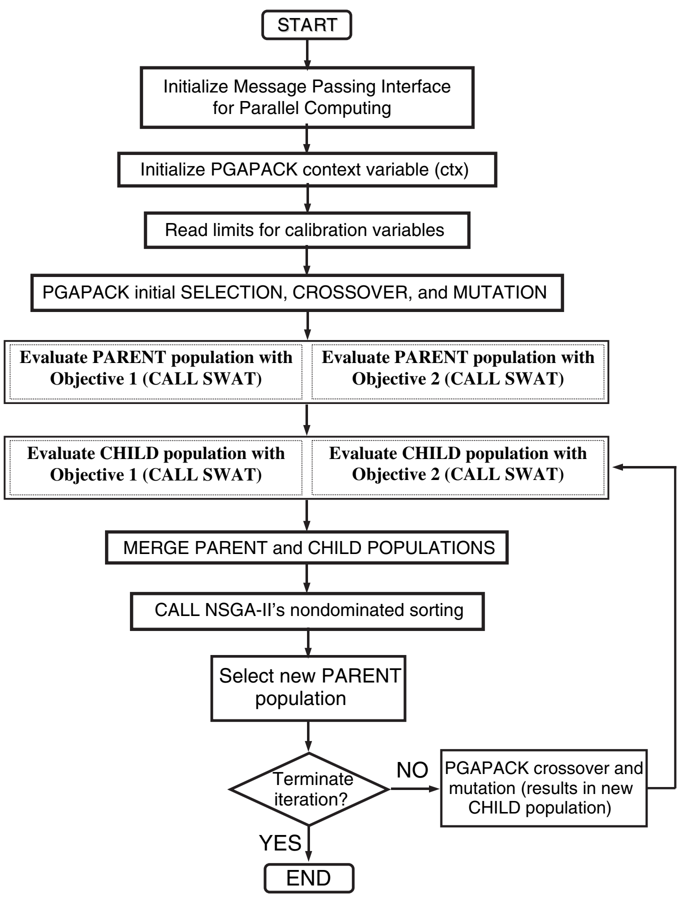

FIGURE 2. Computational Scheme in PGAPACK Linking NSGA-II and SWAT With Two Objective Functions.

The streamflow was designated as driven when the first pass base flow was <80% of the observed streamflow; otherwise the streamflow was classified as nondriven.

The Nash-Sutcliffe model efficiency (Nash and Suttcliffe, 1970) was used to evaluate SWAT's overall performance at calibration and validation:

$$\mathrm{NSE} = \left[1 - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(Q_{\mathrm{obs},i} - Q_{\mathrm{sim},i}\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\left(Q_{\mathrm{obs},i} - \bar{Q}_{\mathrm{sim},i}\right)^2}\right], \tag{2}$$

where $\bar{Q}_{\mathrm{obs}}$ is average of the observed daily flows and all the other variables are as previously defined. The Nash-Sutcliffe efficiency ranges from negative infinity to 1, with 1 indicating a perfect fit.

RESULTS AND DISCUSSION

The optimization run of 1,000 iterations took 10.75 h or an average of 38.7 s per iteration. A single iteration took about 9 min in a P4 3.2 GHz desktop computer. Except for the first iteration, there were 200 function evaluations (two objectives × 100 solutions) and SWAT was executed in each evaluation. Thus, during the entire calibration period there were 200,200 SWAT runs. PGAPACK's evaluation function returns a single scalar value instead of an array. Due to this constraint and other limitations of the PGAPACK library and the parallelization method, a more efficient algorithm that could run SWAT once for each solution per iteration was difficult to implement. This algorithm would have decreased the optimization run by half the current time and is presently explored and developed.

Figure 3 shows the evolution of the Pareto optimal front in the objective space at different generations. Previous optimization runs by the authors showed that there was minimal change in the Pareto optimal front and the objective function values at 1,000th generation; thus, the maximum number of iterations was set to 1,000. The daily Nash-Sutcliffe efficiency values ranged from 0.85 to 0.86 for the final set of solutions. These values were in the upper bounds of the efficiency values (0.65 to 0.86) previously reported for SWAT (Eckhardt and Arnold, 2001; van Griensven and Bauwens, 2003; Eckhardt *et al.*, 2005). All the final solutions were in the first nondominated (most optimal) rank or front (see Deb, 2001; Deb *et al.*, 2002).

Objective 1 and Objective 2 may not strictly conflict with each other but the Pareto optimization resulted in a set of optimal solutions that accounts the trade-off between the objectives as shown in Figure 4. Solution 1 had the lowest RMSE of driven flows but had also the highest RMSE of nondriven flows. In contrast, Solution 2 had the lowest RMSE of nondriven flows but had also the highest RMSE of driven flows. Between these two extreme solutions, there are other solutions with varying degree of tradeoffs between Objective 1 and Objective 2. The choice between the solutions therefore depends on the interest of the user and the selection of the solution for calibration and validation is arbitrary as all the solutions are Pareto-optimal.

The statistical difference between the resulting flows from these solutions was not investigated in this study but focus was on the practical applications of the Pareto front in decision and policy-making. For example, one concerned with sediment concentration would select solutions with lowest RMSE for driven flows as high sediment concentrations are usually associated with high flows. As shown in Figure 5, Solution 1 simulated the streamflow better during an extreme storm event. On the other hand, the solution with lowest RMSE for nondriven flows (Solution 2) is preferred if one is interested in analyzing the base flow (Figure 4). Solution 2 has a better simulation of streamflow during a low flow period without rainfall (Figure 6).

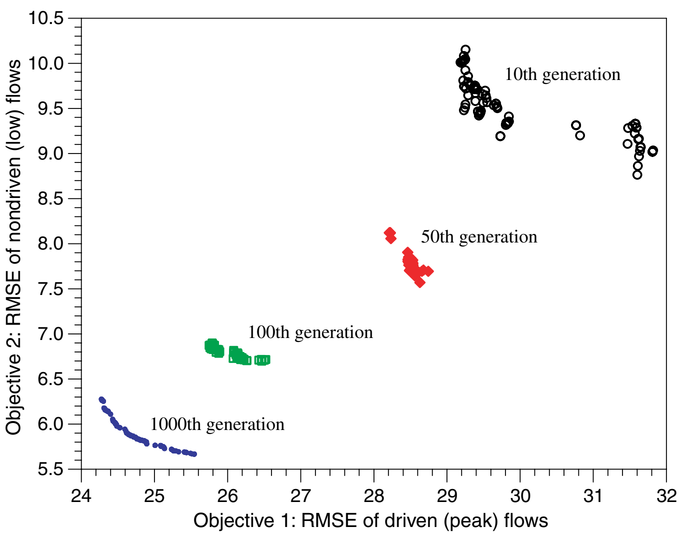

FIGURE 3. Evolution of the Pareto Front in the Objective Space.

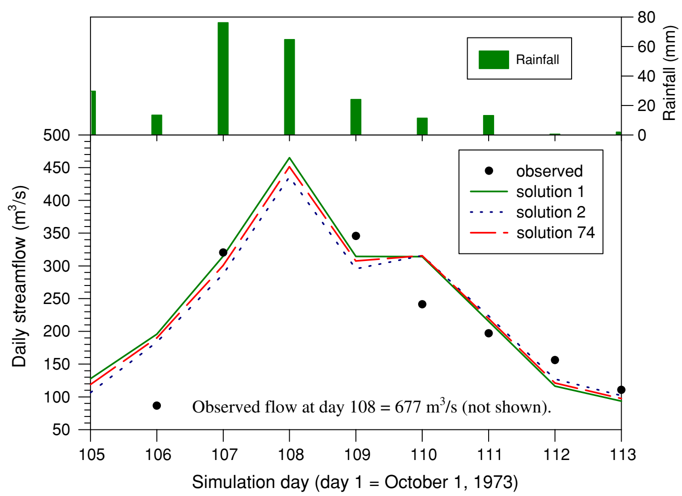

FIGURE 5. Simulated Streamflow of Selected Solutions After 1,000 Generations: Extreme Storm Event.

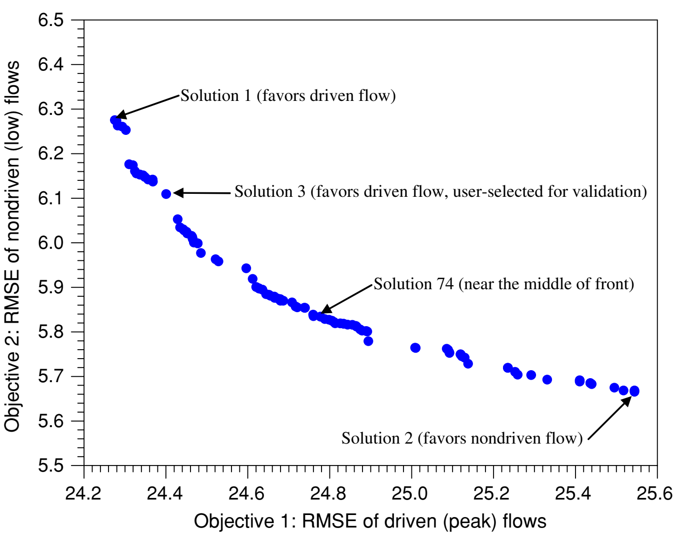

FIGURE 4. Nondominated (Pareto) Front in the Objective Space After 1,000 Generations and Selected Solutions.

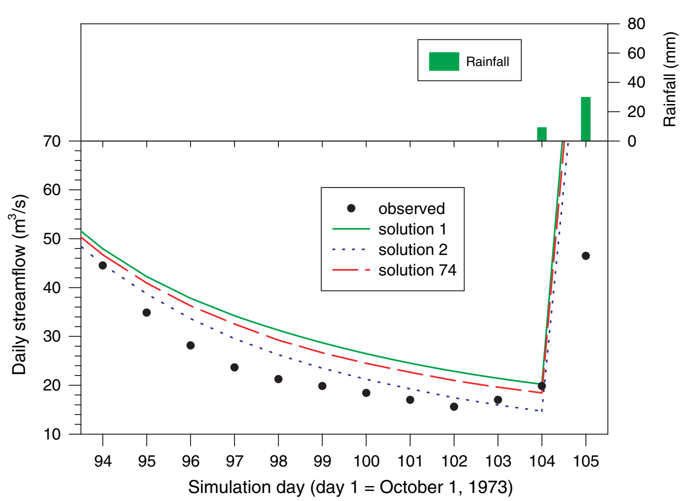

FIGURE 6. Simulated Streamflow of Selected Solutions After 1,000 Generations: Low Flow Period.

The curve number (CN2), soil evaporation compensation factor (ESCO), and available soil water holding capacity (AWHC) govern the surface water response in SWAT (Eckhardt and Arnold, 2001; Van Liew *et al.*, 2005). These parameters favor the contribution of direct runoff to the streamflow. Thus, optimizing Objective 1 (RMSE of driven flows) could in turn optimize these parameters. In the same manner, optimizing Objective 2 (RMSE of nondriven flows) could optimize the parameters that govern the subsurface water response in SWAT. These parameters were the ground-water "revap" coefficient (GWREVAP), ground-water delay (GWDELAY), threshold depth of water in the shallow aquifer for return flow (GWQMN), threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur (REVAPMN), and the base-flow alpha factor (ALPHABF).

The other optimized parameters may not directly influence surface runoff or ground water flow but affect SWAT's routing processes and further adjust the shape of the hydrographs. The surface runoff lag time (SURLAG) governs the release of the surface runoff to the main channel. The effective hydraulic conductivity in tributary channels (CHK1) controls the surface runoff transmission losses from the subbasins to the main channel. The main channel effective hydraulic conductivity (CHK2) shows the relationship of the stream with the ground water and directs the water movement from streambed to the subsurface (or vice-versa) depending on the stream type.

All the 139 parameter values were explicitly optimized in the calibration in contrast to previous studies (Eckhardt and Arnold, 2001; van Griensven and Bauwens, 2003; Eckhardt *et al.*, 2005) where the number of parameters was reduced by either creating sharing parameters between HRUs or fixing ratios between the parameters. However, sensitivity analysis and the modeler's understanding and knowledge of the watershed characteristics are still essential in identifying parameters that need to be optimized.

The calibrated SWAT model simulated well the daily streamflow of the Calapooia River watershed for a 3-year period (October 1973 to September 1976). The selected solution (Solution 3) for calibration resulted in a daily Nash-Sutcliffe efficiency of 0.86, which was a large improvement from 0.28 calculated from the simulated daily streamflow using the default model setup with AVSWAT2000. Figure 7 shows the observed daily streamflow with the lower and upper bounds of the simulated daily streamflow generated from all solutions. As all solutions are optimal, the lower and upper bounds of the simulated flows were comparable with the observed values over the entire simulation period. A high model efficiency of 0.98 was calculated for the 36 monthly means for the same calibration period (Figure 8). Streamflow validation in the following three water years (October 1976 to September 1979) resulted in overall model efficiencies of 0.81 (daily) and 0.95 (monthly), which further verified the calibration results (Figure 9).

Despite the high daily Nash-Sutcliffe efficiency coefficients, the simulation outputs tend to underestimate high peak flows. Previous studies (Eckhardt and Arnold, 2001; Van Liew *et al.*, 2005) reported large differences between the observed and simulated values at streamflow peaks, despite the optimization of parameters that favor contribution of direct runoff to streamflow. It might be possible

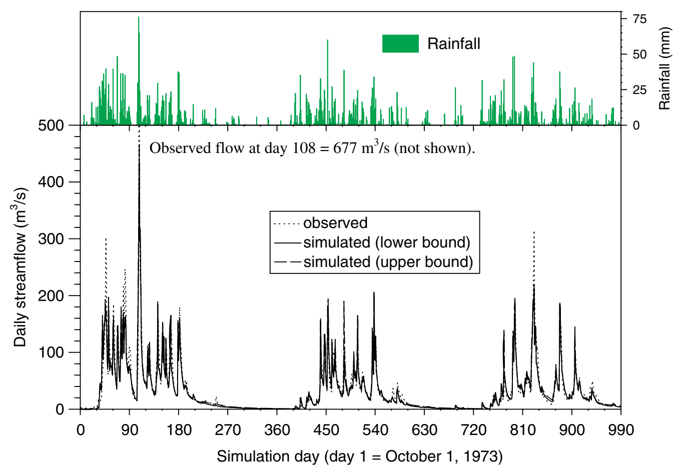

FIGURE 7. Calibration Results: Observed and Simulated Daily Flows for the Calapooia Watershed.

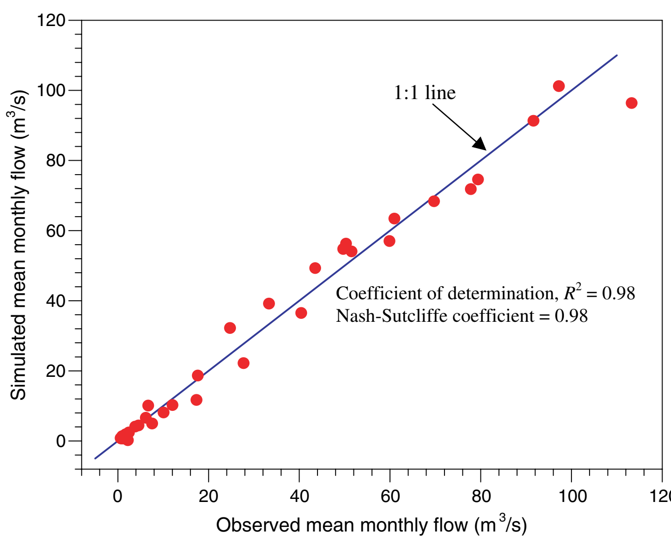

FIGURE 8. Calibration Results: Observed and Simulated Monthly Flows for the Calapooia River Watershed Fitted in a 1:1 Line.

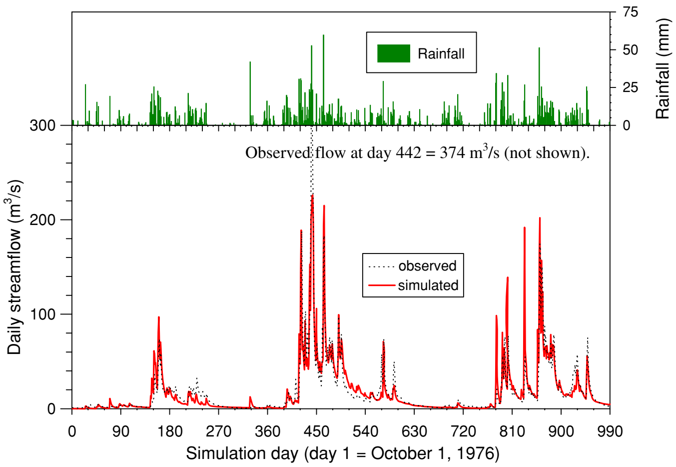

FIGURE 9. Validation Results: Observed and Simulated Daily Flows for the Calapooia Watershed.

that the soil properties were not attributed to the correct hydrologic soil groups, and importance were given to land covers and soils dominating the HRUs (Eckhardt and Arnold, 2001). These uncertainties might have been reduced with the use of the county wide and much detailed soil survey geographic (SSURGO) database.

This underestimation could also be the effect of using RMSE as objective functions in the optimization. The RMSE is a function of the square of the difference between the observed and simulated streamflow; which makes the optimization highly sensitive to extreme peak flows. Boyle *et al.* (2000) reported that the use of RMSE as a criterion in automatic calibration leads to strongly biased simulations of the recessions. It was not surprising then that the Nash-Sutcliffe efficiencies for the driven flows ranged from 0.83 to 0.84 and higher efficiencies for nondriven flows from 0.92 to 0.93.

SUMMARY AND CONCLUSIONS

This study investigated the application of MOEA and Pareto ordering optimization in the automatic calibration of the Soil and Water Assessment Tool (SWAT), a complex, process-based, semi-distributed, and continuous hydrologic model. SWAT was automatically calibrated for the Calapooia watershed in Oregon, USA, by optimizing selected parameters with two objective functions. The NSGA-II, a fast MOEA, and SWAT was called in FORTRAN from a PGAPACK to determine the Pareto optimal set.

The automatic multi-objective calibration of a complex process-based watershed model was successfully implemented using Pareto ordering optimization and MOEA. Pareto ordering optimization gives the modeler a set of optimal solutions that accounts the trade-offs between the objectives. The selection of a solution depends on the modeler's preference and interest as all the solutions are Pareto optimal. The calibrated SWAT model simulated well the daily streamflow of the Calapooia watershed for a 3-year period. The daily Nash-Sutcliffe coefficients were 0.86 at calibration and 0.81 at validation. Future studies include: (1) simultaneous automatic calibration of water quality and quantity parameters, and (2) link with economic models. The application of Pareto optimization in decision and policy-making problems related to conflicting objectives of economic costs and environmental quality will also be explored.

ACKNOWLEDGMENT

This study was funded by USDA-ARS through the Conservation Effects Assessment Project (CEAP).

LITERATURE CITED

Arnold, J.G. and P.M. Allen, 1999. Automated Methods for Estimating Baseflow and Ground Water Recharge From Streamflow Records. *Journal of the American Water Resources Association* 35(2):411-424.

Arnold, J.G., P.M. Allen, R. Muttiah, and G. Bernhardt, 1995. Automated Base Flow Separation and Recession Analysis Techniques. *Ground Water* 33(6):1010-1018.

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams, 1998. Large Area Hydrologic Modeling and Assessment. Part I: Model Development. *Journal of the American Water Resources Association* 34(1):73-89.

Boyle, D.P., H.V. Gupta, and S. Sorooshian, 2000. Toward Improved Calibration of Hydrologic Models: Combining Strengths of Manual and Automatic Methods. *Water Resources Research* 36(12):3663-3674.

Boyle, D.P., H.V. Gupta, S. Sorooshian, V. Koren, Z. Zhang, and M. Smith, 2001. Toward Improved Streamflow Forecasts: Value of Semidistributed Modeling. *Water Resources Research* 37(11): 2749-2759.

Deb, K., 2001. *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley and Sons, West Sussex, UK, 515 pp.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan, 2002. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2):182-197.

Di Luzio, M., R. Srinivasan, and J.G. Arnold, 2001. *Arcview Interface for SWAT2000: User's Guide*. Agricultural Research Service, USDA, Temple, Texas. http://www.brc.tamus.edu/swat/downloads/doc/swat2000.pdf, *accessed* October 2006.

Duan, Q., S. Sorooshian, and H.V. Gupta, 1992. Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models. *Water Resources Research* 28(4):1015-1031.

Eckhardt, K. and J.G. Arnold, 2001. Automatic Calibration of a Distributed Catchment Model. *Journal of Hydrology* 251(1-2):103-109.

Eckhardt, K., N. Fohrer, and H.G. Frede, 2005. Automatic Model Calibration. *Hydrological Processes* 19(3):651-658.

Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Longman, Inc., New York, 412 pp.

Gupta, H.V., S. Sorooshian, and P.O. Yapo, 1998. Toward Improved Calibration of Hydrologic Models: Multiple and Noncommensurable Measures of Information. *Water Resources Research* 34(4):751-763.

Holland, J.H., 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, Michigan, 183 pp.

Khu, S.T. and H. Madsen, 2005. Multiobjective Calibration With Pareto Preference Ordering: An Application to Rainfall-Runoff Model Calibration. *Water Resources Research* 41(3):W03004.

Levine, D., 1996. *Users Guide to the PGAPack Parallel Genetic Algorithm Library, ANL-95/18*. Argonne National Laboratory, Argonne, Illinois. http://www-fp.mcs.anl.gov/CCST/research/reports_pre1998/comp_bio/stalk/pgapack.html, *accessed* October 2006.

Nash, J.E. and J.E. Suttcliffe, 1970. River Flow Forecasting Through Conceptual Models – Part I: A Discussion of Principles. *Journal of Hydrology* 10(3):282-290.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, J.R. Williams, and K.W. King, 2002. *Soil and Water Assessment Tool User's Manual: Version 2000*. Agricultural Research Service, USDA, Temple, Texas. http://www.brc.tamus.edu/swat/downloads/doc/swat2000theory.pdf, *accessed* October, 2006.

van Griensven, A. and W. Bauwens, 2003. Multiobjective Autocalibration for Semidistributed Water Quality Models. *Water Resources Research* 39(12):1348.

Van Liew, M.W., J.G. Arnold, and D.D. Bosch, 2005. Problems and Potential of Autocalibrating a Hydrological Model. *Transactions of the ASAE* 48(3):1025-1040.

Vrugt, J.A., H.V. Gupta, L.A. Bastidas, W. Bouten, and S. Sorooshian, 2003. Effective and Efficient Algorithm for Multiobjective Optimization of Hydrologic Models. *Water Resources Research* 39(8):1214.

Whittaker, G., 2004. Use of a Beowulf Cluster for Estimation of Risk Using SWAT. *Agronomy Journal* 96(5):1495-1497.

Yapo, P.O., H.V. Gupta, and S. Sorooshian, 1998. Multi-Objective Optimization for Hydrologic Models. *Journal of Hydrology* 204(1):83-97.